DOCUMENT RESUME

ED 230 607 TM 330 402

AUTHOR Brinberg, David; McGrath, Joseph E.
TITLE A Validity Network Schema.
PUB DATE Apr 83
NOTE 33p.; Paper presented at the Annual Meeting of the American Educational Research Association (67th, Montreal, Quebec, April 11-15, 1983).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Behavioral Science Research; Generalization; Measurement Techniques; Models; *Research Methodology; *Research Needs; Research Problems; *Research Tools; *Validity
IDENTIFIERS Robustness; *Validity Network Schema; *Validity Research

ABSTRACT

Numerous forms and meanings of validity have been developed to provide researchers with the opportunity to assess the many potential sources of ambiguity that exist in any research finding. A Validity Network Schema (VNS) developed by Brinberg & McGrath (1982) is extended and elaborated in order to describe the components of the research process and their interrelations, and draw out some implications of the schema for the acquisition of knowledge. VNS describes research as a three-stage process, whose outcome is some structural combination of concepts, methods, and substantive events. In stage 1, researchers develop, clarify, and select elements and relations from the three basic domains--conceptualization, methodological, and substantive. Stage 2, the research study proper, involves two steps: the researcher combines elements and relations from two of the three domains to form an intermediate structure; and the researcher integrates that structure with the elements and relations from the third domain. Stage 3 of the research process involves the reduction of uncertainty associated with the findings of Stage 2 by exploring the range of respondents, events, contexts, methods, and concepts over which the stage 2 results do and do not hold. In stage 3, validity takes on the meaning of robustness and generalizability. (PN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED230607

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

A Validity Network Schema

David Brinberg & Joseph E. McGrath

University of Maryland

University of Illinois

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D. Brinberg

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper to be presented at American Educational Researchers Association, April. 1983.

TM830402

Abstract

The concept of validity has received a great deal of attention in the behavioral sciences. Numerous forms and meanings of validity have been developed to provide researchers with the opportunity to assess the many potential sources of ambiguity that exist in any research finding. A Validity Network Schema developed by Brinberg & McGrath (1982) is extended and elaborated in order to: (a) describe the components of the research process and their interrelations, and (b) draw out some implications of that schema for the acquisition of knowledge. A detailed discussion is presented on forms of uncertainty that influence what researchers generally describe as external validity.

The quest for validity is one way to describe systematic efforts to build research information and to establish confidence in it. In other words, to pursue validity is to try to reduce the uncertainty associated with a set of research findings. Considering the relatively strong emphasis that psychology and other behavioral sciences have given to empirical data, and how to acquire and process such data, those fields have given relatively little attention to understanding the processes by which research knowledge is acquired -- that is, to how we gain confidence that we "know" what the data we have gathered means.

We have had some help from philosophers of science (e.g., Ayer, 1959; Popper, 1959), who have clarified the epistemological bases for our research activities (e.g., logical positivism, the falsification approach). There also has been considerable discussion among behavioral scientists themselves, much of it relatively recently, on the methodological underpinnings of our work. This has included treatises on the nature and components of the research process; classification of types of research (e.g., lab vs. field vs. survey; basic vs. applied) and of their fundamental differences; and examination of conditions that threaten or limit the potential usefulness of a set of research findings. (See for example, Blalock, 1982; Campbell & Fiske, 1959; Campbell & Stanley, 1966; Cook & Campbell, 1979; Cronbach, 1982; Kaplan, 1964; McGrath, Martin & Kulka, 1982; Runkel & McGrath, 1972; Sjoberg & Nett, 1968; Wimsatt, 1981).

Falsification, Triangulation, and Differentiation

The works cited above share a number of underlying themes. One such theme underlying much of that work is Popper's (1959) view of how we acquire knowledge. In Popper's (1959) falsification approach, researchers

acquire knowledge either: (a) by failing to disconfirm some expectation or hypothesis; or (b) by confidently disconfirming such an hypothesis.

Another theme underlying these works is the idea of triangulation, or more generally, the idea of multiple operationalism. That idea is, we believe, an integral part of the Popperian position and much of the other methodological and epistemological work cited previously (e.g., Wimsatt, 1981; McGrath, Martin & Kulka, 1982; Runkel & McGrath, 1972). At its core, the triangulation position holds that we gain confidence in a research finding (i.e., we eliminate threats to the validity of that finding and reduce the uncertainty vis a vis that finding), only when we have convergence of substantive outcomes based on the use of different and independent models, methods, and occasions. Campbell and Fiske (1959) present the idea of triangulation when they discuss the convergent and discriminant validity of measures. Wimsatt (1981) also uses the triangulation ideas when discussing robustness analysis of research findings:

> "1. To analyze a variety of independent derivation, identification or measurement processes; 2. to look for and analyze things which are invariant over or identical in the conclusions or the results of these processes; 3. to determine the scope of the processes across which they (the findings) are invariant and the conditions on which this invariance depends; and 4. to analyze and explain any relevant failures of invariance." (p. 126).

Wimsatt's (1981) idea of "failures of invariance" makes the very important point that the researcher is not only looking for the conditions under which the finding will fit the hypothesis, but also should be trying to identify and explain the conditions under which the findings are inconsistent with (i.e., disconfirm) the hypothesis. Lynch (1982) has talked about such non-confirmatory outcomes, such "failures of invariance," as the lack of construct validity. Our position is that both "invariance" and

"failures of invariance" can yield research information by identifying the scope and the limits of a set of research findings.

To anticipate some terminology we will use in later sections of this paper, the research process involves elements and relations from each of three domains: conceptual, methodological, and substantive. The need for triangulation is an important feature of many aspects of the research process and applies in each of those three domains.

In the conceptual domain, for example, Feyerabend (1970) draws upon the idea of triangulation when he discusses the need for researchers to develop competing conceptual models and argues that theoretical pluralism is assumed to be an essential feature of all knowledge that claims to be objective. Similarly, Garner, Hake, & Erikson (1956) apply the principle of triangulation within the conceptual domain when they argue that a research discipline is likely to advance only when it compares alternative theories.

The idea of triangulation is most often associated with the methodological domain. Campbell (1981) and colleagues have argued for triangulation across methods and regard triangulation as the method of choice for reducing certain sources of potential invalidity (i.e., for reducing uncertainty) in the measurement and manipulation of variables. Both Fiske (1982) and Campbell & O'Connell (1982) offer recent reviews of some work in that tradition. While these works are essentially concerned with the validity of measures, the triangulation approach also has been advocated with regard to validity of relations between measures. McGrath, Martin & Kulka (1982), for example, argue that successive studies of the same problem should make use of maximally different research strategies (e.g., laboratory experiments, field studies, sample surveys) and different

research designs. Such divergence of methods will increase our confidence in our findings (if the findings converge), because it will let us offset the weaknesses of any one strategy or design by using other strategies or designs that are equally, though differently, flawed.

The triangulation idea also applies in the substantive domain. The behavioral science community has long recognized that any interpretation of findings must be considered in terms of the limits of the samples (of subjects, events, and contexts) on which any set of studies is based. For example, some behavioral scientists have long been concerned that psychology relies too much on evidence gained from research on college students under artificial conditions. Such a concern implies a call for triangulation of research findings across different facets of the substantive domain (i.e., across subjects, events, and contexts). But the need for such exploration of the scope and limits of findings, with respect to facets of the substantive domain, is too often honored in the breech rather than in the observance. At times, the standard caveat that we insert near the end of our reports of empirical investigations (e.g., "Future research is needed using other samples and stimulus materials, to determine the robustness...") is juxtaposed with conclusions that violate that caveat.

Triangulation has a complement; the idea of differentiation. The dual concepts of triangulation and differentiation -- of scope and limits -- are expressed by Wimsatt (1981) as invariance and failures of invariance. Both are useful. Both are necessary strategies for the acquisition of knowledge. Both apply in each of the three domains. In the conceptual domain, we are concerned with determining whether some model is sufficient and adequate to account for the set of findings when compared to alterna-

native models. In the methodological domain, we seek both convergence and discrimination of our measures, our research designs, and our research strategies. In the substantive domain, too, we need to establish the range (of subjects, events, and contexts) over which our findings do and do not hold.

These dual concepts -- triangulation and differentiation -- are directly connected to the falsification principles. Triangulation is the acquisition of knowledge by the first falsification strategy; that is, failing to disconfirm that two things (e.g., two measures) are the same. Differentiation is the acquisition of knowledge by the second falsification strategy; that is, confidently disconfirming that two things are the same. To focus on only one of these strategies is to restrict our ability to reduce uncertainty in any set of research findings. In other words, it limits our ability to know what we know.

Recently, Brinberg and McGrath (1982) presented a conceptual framework called a Validity Network Schema. That presentation defines the components of the research process, lays out the interrelations among those components, traces certain characteristic patterns in the conduct of research, and draws out some implications of that view for various forms and meanings of validity. We will draw heavily upon that schema, here, placing special emphasis on forms of uncertainty that influence what researchers generally describe as external validity.

Validity Network Schema

Overview of Research Process

All types of research involve the combination of some set of concepts, some set of methods for making observations or comparing sets of observations, and some set of substantive events and phenomena that are to be the focus of the study. The validity network schema proposed by

Brinberg & McGrath (1982) describes the research process as the selection, combination and use of elements and relations from the conceptual, the methodological, and the substantive domains.

Brinberg & McGrath (1982) define the conceptual domain as containing concepts (element level) and conceptual models (relation level). The methodological domain contains instruments and techniques for making observations (element level) and structures or models for comparing and contrasting sets of observations (relation level). The substantive domain contains events (element level) and phenomena or processes (relation level).

The research process has three stages. Stage 1 involves development, clarification, and selection of elements and relations in each of three domains. Stage 1 is preparatory for later stages. Stage 2 involves two steps, and is usually what we consider a research study. In step 1, the researcher combines elements and relations from two of the domains to form an intermediate structure; in step 2, the researcher integrates the structure with elements and relations from the third domain. With three domains, there are three paths by which this two step process can be carried out. Stage 3 involves replicating the activities of Stage 2 as well as selecting elements and relations from each of the domains that are considered to be similar or different than previous work in order to estimate the scope and limits of Stage 2 results.

Associated with each stage of the research process is a different meaning of validity. In Stage 1, validity takes on the meaning of value. Elements and relations are developed, clarified and selected from each of the three domains if a researcher[1] considers them to be "of value" -- interesting or useful. In Stage 2, validity takes on the meaning of correspondence or fit; that is, validity is construed as the extent to which elements and relations from the different domains fit when paired together. In Stage 3, validity takes on the meaning of robustness; that is, validity is increased to the extent that Stage 3 activities increase

our confidence (i.e., reduce our uncertainty) concerning a (Stage 2) research finding. These concepts are summarized in graphic form in Figure 1. Because the primary aim of this paper is to elaborate on the validity issues associated with the third stage of the research process, i.e., issues typically associated with external validity, the role of values in the research process, and of forms of validity considered to be part of a research study proper (i.e., the internal validities) will not be presented. The reader interested in a more detailed discussion of the Stage 1 and Stage 2 validity issues should read Brinberg & McGrath (1982).

Uncertainty Reduction: Stage 3 - Validity as Robustness

As noted in an earlier section, researchers acquire knowledge by: (a) failing to disconfirm some hypothesis or (b) being confident some hypothesis has been disconfirmed. Both these forms of acquisition of knowledge are used to reduce the uncertainty associated with a research finding (from Stage 2). In Stage 3, a researcher can reduce uncertainty by addressing three issues. The first issue is: If the study were repeated, exactly, would the same findings occur? That is, would the study replicate. Within the Validity Network Schema, replication would imply that a researcher had selected concepts, methods, and substantive events that were assumed to be the same as the previous set and expected the findings to be the same as well. At an element level, researchers typically refer to this validity issue as the question of the reliability of a measure. At the relations level, the validity issue involved in replication is part of what Cook & Campbell (1979) call statistical conclusion validity.

The second issue of Stage 3 is: If the study were repeated and the researcher (a) selected elements or relations that were different from previous work with respect to some facet(s) of one of the domains and (b)

made those selections such that the findings were expected to be similar to those of the prior study, would the findings in fact turn out to be robust across these differences? Since there are three domains, there are three distinct areas of robustness that need to be examined.

The third issue to be addressed in Stage 3 is: Under what conditions will the findings not hold? In other words, what are the boundaries (limits) across which the findings do not hold? As for the robustness question, since there three domains, there are three sets of boundaries that need to be determined -- conceptual, methodological, and substantive.

Both the effort to replicate a finding as well as conduct a robustness analysis are attempts to reduce uncertainty by failing to disconfirm some hypothesis. When replicating a study, as well as conducting a robustness analysis, the desired (and anticipated) outcome is a failure to disconfirm that the original study and the replicate have similar findings. (i.e., find no differences). When a researcher does not find differences, the uncertainty associated with the prior finding is reduced. When a replication is attempted and differences do occur between the original and subsequent study, however, there is increased uncertainty associated with the prior finding. On the other hand, when a robustness analysis finds differences (and when the researcher can be confident in the finding), this identifies one boundary of the prior finding. In other words, finding differences also can contribute to a reduction of uncertainty because it identifies conditions under which the findings do not hold. A boundary search study is an attempt to determine those conditions under which a finding does not hold; that is, it is an attempt to acquire knowledge by confidently disconfirming a prior finding.

These three approaches for reducing uncertainty about a research finding -- replication, robustness analysis and boundary search -- use the two falsificationist strategies for acquiring knowledge: (1) failing to disconfirm some prior finding and (2) confidently disconfirming some prior finding.

Typically, researchers focus their Stage 3 efforts on the first falsificationist strategy by attempting to replicate research findings and, perhaps, by determining their robustness with respect to one facet (e.g., respondents) of one domain (e.g., substantive). But, researchers seldom deliberately examine the boundaries associated with their findings, that is, they seldom use the second falsificationist strategy in their Stage 3 search for uncertainty reduction. When outcomes of such a study lead to non-confirmation of the robustness of a set of findings, these are usually regarded, pejoratively, as "negative results" -- that is, failures of invariance in Wimsatt's terms. Our contention is that these so-called "negative results" can provide researchers with information just as useful as would "positive results" because those negative results also can reduce the uncertainty associated with a research finding. A more detailed discussion of these issues will be presented in later sections.

As noted earlier, there are three domains with respect to which a researcher needs to reduce uncertainty -- conceptual, methodological and substantive. The replication question is not specialized by domain, nor is it particular to any given facet of any domain. When a replication is conducted, a researcher tries to reproduce the conditions of a previous study in all its particulars. In fact, replication can be considered a special case of robustness analysis. In the following parts of this section, therefore, we will examine aspects of uncertainty reduction associated with the other two questions of Stage 3: robustness analysis and boundary search for each of the three domains. Table 1 contains a summary of the facets to be examined in each of the domains.

Insert Table 1 here

Substantive Domain - Searching for ecological validity

Robustness analysis in the substantive domain has traditionally been referred to as external validity. Generally, researchers will explore the robustness (or external validity) of their findings by assessing the extent to which a particular research finding will "generalize across" (Cook & Campbell, 1979) some designated population. Much of the literature on the external validity of a research finding has focused on one of the three aspects typically examined in the substantive domain; that is, the type of subject used in the research studies (e.g., college students vs. "real" people). Sometimes this debate seems to hinge on the question of which type of sample is the proper kind on which research findings are to be built. What seems more cogent, to us, is not which type of group is used in any one study, but rather that different samples are used in different studies so that we can examine the robustness of research findings across variations in that facet. When findings gathered on different types of respondents converge, the uncertainty associated with those research findings is reduced.

A second facet within the substantive domain on which robustness is frequently assessed is variations across sets of events or behaviors. Generally, researchers are interested in explaining some range of events. To do this, the researcher needs to examine the full range of events that are purported to be explained by sampling different sets of events from the substantive domain and determining the similarity of findings across the different studies. For instance, theories of decision making (e.g., Fishbein & Ajzen, 1975; Jaccard, Knox & Brinberg, 1979) purport to explain

a wide range of decisions. The robustness of each theory can be assessed by examining different classes of decisions. The uncertainty associated with a particular theory is reduced to the extent that similar findings occur across studies involving different classes of decisions.

A third facet within the substantive domain about which robustness is frequently explored concerns the spatial and temporal context surrounding the event. Researchers examining the person by situation interaction (e.g., Endler & Magnussen, 1976) suggest that the context within which an event occurs modifies its meaning. To reduce the uncertainty associated with a finding, the researcher can assess its robustness by sampling events and phenomena from the substantive domain such that they vary systematically with respect to the context in which the events occur.

In addition, the temporal context in which an event or phenomenon is observed may influence the empirical findings. For instance, researchers (e.g., Kahle & Berman, 1979) using cross-lagged analysis techniques to study the relation between attitudes and behaviors need to be concerned about the robustness of their findings across different time periods. If a similar causal direction of a relation (e.g., attitude X causes behavior Y) is found for a wide variety time periods, the researchers will have reduced some uncertainty concerning the nature of the empirical finding.

To summarize, research exploring the robustness of findings across the elements (i.e., events) and relations (i.e. phenomenon) from the substantive domain has focused mainly on three facets: (a) the sample of respondents in the study, (b) the range of events/phenomena to which the finding is applied and (c) the physical and temporal context surrounding the event/phenomenon. These are only three of the many facets relevant to ecological validity.

Boundary search is an alternative to robustness analysis and also can be used to reduce the uncertainty associated with a set of findings i.e., when the researcher is able to confidently disconfirm a finding by testing it under conditions for which it is expected not to hold. Runkel & McGrath (1972) point out that knowledge is a knowledge of differences. A finding that is unbounded provides no useful information since no distinctions are made. For a finding to be useful, a researcher needs to identify not only its scope but also its limits. Lynch (1982) discusses such boundary search as attempts to identify background factors that interact with a finding. By incorporating these background factors into a more complex statement of the empirical finding, a researcher reduces uncertainty by articulating both the scope and boundaries, (i.e., those conditions under which a finding should or should not occur). Only when all boundaries have been specified can a researcher be completely confident in the knowledge acquired (that is, have reduced uncertainty completely).

Regarding respondents, a researcher can try to identify those types of respondents for whom a particular finding does not hold. For instance, Jaccard, Knox & Brinberg (1979) used a subjective probability model to predict voting behavior of both highly educated and poorly educated respondents. The model predicted the behavior of the highly educated respondents more effectively than it did for the poorly educated group. Those authors, thereby, have identified one limit or boundary of the findings of their model, which suggests that future work should consider the influence of education level and related attributes on the models predictions.

Similarly, the researcher can try to identify the range of events (behaviors) over which a research finding does or does not hold. Consumer

behavior researchers, for instance, (e.g., Engel & Blackwell, 1982) have distinguished between low-involvement and high-involvement decisions. Apparently, different decision processes occur for these two types of decisions. Given this information, a researcher can now specify the type of decisions that can and cannot be accurately predicted from a given model; thus, reducing the uncertainty associated with the findings.

The researcher also needs to determine the boundaries of the spatial and temporal context beyond which a finding will not hold. Cronbach (1975) discussed such a boundary search of the context facet when he illustrated the influence of higher-order interactions. Specifically, he analyzed a situation in which a certain dose of a sedative was effective for 30 minutes in one context (i.e., certain type of wood chip was used for the animal's bedding) and only 15 minutes in another context (i.e., a different set of wood chips). Identifying such an interaction amounts to specifying a context boundary and, thereby, serves to reduce the uncertainty associated with the effect of drug dosage.

In the substantive domain, knowledge can be acquired by identifying the boundaries associated with variations along each of many facets (e.g., respondents, events and contexts). Simply determining the robustness or scope of a finding is not sufficient; the limits of the finding also must be identified. The ecological validity of a finding can be assessed by conducting robustness analysis and boundary search; that is, by establishing the scope and limits of the findings.

Searching for Validity in the Methodological Domain.

In the previous discussion of ecological validity, we assumed that both the methods and the concepts associated with a finding were held constant and that only elements and relations selected from the substantive

domain were allowed to vary. But in addition to determining the robustness and boundaries of findings with respect to the substantive domain, a researcher needs to assess the robustness and boundaries (the scope and limits) of the finding with respect to various facets of the methodological domain.

Robustness analysis in the methodological domain has traditionally been thought of as assessing "method variance." When examining this issue, researchers typically select different measurement techniques (e.g., different tests of the same trait) and use some variant of the multitrait-multimethod paradigm to determine the extent of method variance. That is, they attempt to assess the robustness of a research finding across different measures intended to measure the same constructs.

Robustness with regard to several aspects of measures need to be considered. One is obtrusiveness, i.e., the extent to which a measure interacts with an event. Another is the set of assumptions inherent in a measure. For instance, when using a non-disguised, self-report measure (e.g., a semantic differential scale), a researcher assumes the subject is both able and willing to report some cognitive or affective state. Several researchers, however, have questioned the ability of respondents to report cognitive states (e.g., Nisbett & Wilson, 1977) and whether a subject's response is equivalent to the psychological representation of this response (Anderson, 1974). Furthermore, subjects may be unwilling to provide an accurate response if they feel it is a socially undesirable (e.g., Crowne & Marlowe, 1964) or if they feel it would put them in an unfavorable light (e.g., Rosenberg, 1969).

The robustness of a research finding with regard to the features of measures can be assessed by comparing a set of findings that differ in

those features. For instance, if researchers are concerned that respondents may give socially desirable responses, they may want to use both a non-disguised (e.g., semantic differential) and a disguised (e.g., bogus pipeline) measurement method. If the different methods yield similar findings, the uncertainty associated with the finding is reduced.

Researchers also need to consider robustness with respect to research designs and research strategies. Runkel & McGrath (1972) describe eight classes of strategies that vary in terms of the degree to which they offer a mix of realism, precision and generalizability. If the evidence on a research problem all comes from only one strategy, there is uncertainty concerning the extent to which the findings are artifacts or features of that strategy. This uncertainty can be reduced if researchers use different strategies and their findings converge across those studies. As with the debate over what type of respondents would be "best" for research, there has been much debate as to the relative merits of laboratory vs. field studies. We would argue, instead, that no strategy is flawless and only through the use of multiple strategies will the uncertainty associated with a set of findings be reduced.

Robustness analysis also may be conducted with regard to research designs. For instance, between-subject vs. within-subject factorial designs may result in different findings because the two types of designs involve different assumptions. Greenwald (1976) has discussed the assumptions associated with these two types of designs and how these may influence results. For example, the findings associated with the just world (Lerner & Miller, 1978) have typically been examined by using a between-subject factorial design. The standard effect (i.e., denigration of positively evaluated victim) may simply be an artifact of the between-

subject design because the respondent does not view all levels of the victim factor. To avoid this potential confound, a researcher could explore the robustness of the finding across different types of designs.

To summarize, research exploring the robustness of a set of findings across the elements (i.e., measures) and relations (i.e., designs, strategies) from the methodological domain has focused mainly on three facets: (a) the measures, (b) the research strategy and (c) the research design. These are only three of the many facets relevant to methodological validity.

The other side of the robustness question here, as in the substantive domain, is the attempt to identify the range of measures, strategies, and designs beyond which a finding will not hold. Often, such information is regarded as "negative information." We would argue, rather, that information about limits is as "positive" as information about scope, in the sense that it helps us reduce uncertainty about the research finding being examined.

For the facet of measures (i.e., elements from the methodological domain), a researcher can acquire knowledge by identifying those classes of measures over which a finding does not hold. For instance, suppose a finding does not hold for both disguised and non-disguised measurement techniques. One possible reaction is to point out the limitations of the particular method for which the finding did not hold. An alternative reaction would be to explore what basic features differentiate these two measurement instruments (e.g., the social desirability of the response) and to incorporate that information as part of the (now elaborated) research

finding. To the extent that a researcher is confident with regard to the properties of the measures, differences in findings that do occur may provide insights concerning the limits or boundaries of that finding. To treat differences in findings merely as artifacts of the different measures is to ignore potentially useful (i.e., uncertainty reducing) information.

The same point of view concerning the boundary search of measures also applies to research strategies. For instance, if a finding from the lab does not replicate in a field setting, a researcher has acquired useful information about the limits (boundaries) of the findings. A good example of the use of this type of information is Hovland and his associates (1949) discusion of lab vs. field results for attitude change research.

For the facet of research design, uncertainty can be reduced when a researcher is able to specify the types of designs that will lead to a particular pattern of findings. If the findings from two different types of designs (e.g., between vs. within subject factorial design) are predictably different, the researcher can incorporate this as useful information that will modify the finding (e.g., carry-over effects, range effects). For instance, Grice & Hunter (1964) showed that quite different results were obtained in simple reaction time data for eyelid conditioning when within subject and between subject designs were used. These results specify one of the limits of the finding; thus, reducing uncertainty.

For these facets of the methodological domain -- measures, strategies, and research designs -- uncertainty is reduced when the researcher is able to specify the conditions under which a finding would not hold. Treating differences in the findings using different methods as

a methodological artifact ignores information that may potentially reduce uncertainty.

The methodological validity of a set of findings can be assessed by conducting robustness analyses and boundary search; that is, by establishing the scope and limits of the findings. This can be accomplished using many aspects of the methodological domain, including measures, designs, and strategies.

Searching for Explanatory Validity in the Conceptual Domain

As researchers, we are interested in conceptual models that uniquely and adequately account for some set of findings. A researcher needs to assess whether other concepts or models could account for the empirical findings equally or more effectively than do the original ones, and to explore what functional form of the relations (e.g., linear, non-linear, asymmetric, recursive) best describes the relations among the concepts.

The researcher can attack these questions by comparing the robustness and boundaries of different elements and relations selected from the conceptual domain. Specifically, a robustness analysis and boundary search will help indicate the extent to which the original model accurately and uniquely accounts for a set of findings, compared to alternative concepts and models, and the range of alternative concepts and models that would account for these findings equally well. Such an approach is consistent with Feyerabend's (1970) discussion of "theoretical pluralism;" that is, of the need to assess the meaningfulness of a model by competitively testing it against alternative models. In the case of robustness analysis in the conceptual domain, the researcher often wishes for the outcome to be a confident disconfirmation of the hypothesis that the original model is no more accurate than the alternatives.

Two facets are especially pertinent to assess the robustness of a set of findings with respect to the conceptual domain. One has to do with the form of the functional relations (linear, recursive, etc); the second has to do with whether the conceptual model is sufficient to account for the findings.

The robustness of the functional relations among a set of concepts may be assessed by contrasting the original relations (i.e., the model) with some specified set of alternatives. For instance, Campbell (1963) presented a model using a threshold function to relate attitudes and behavior, as an alternative to the linear relation typically tested. To the extent that the linear relation more accurately describes the attitude-behavior relationship, the threshold function may be rejected as an alternative model -- or vice versa. Such an analysis reduces the uncertainty concerning the relationship between attitudes and behavior to the extent that it lets us eliminate one set of potential alternatives.

The sufficiency of a conceptual model may be assessed by comparing the original model with alternatives. Brinberg (1979) described this approach when comparing two models of decision making. Specifically, he compared the sufficiency of the Fishbein model relative to the Triandis model of decision making. In the prediction of behavior, Fishbein's model was found to be robust; that is, alternative formulations for predicting behavior were less accurate than the Fishbein model. In another research area, Birnbaum and Mellers (1979) postulated and found that a single factor model of stimulus recognition described more accurately the relations between affect and exposure than a two-factor model. In both examples, uncertainty was reduced because alternative models did not account for the findings as accurately as the original model.

To summarize, the robustness of a research finding across the elements (i.e., concepts) and relations (i.e., conceptual models) selected from the conceptual domain has focused mainly on two facets: (a) the functional relations among the elements and (b) the sufficiency of the conceptual model. The analysis of these facets is one aspect of explanatory validity, i.e., the extent to which the uncertainty concerning the elements and relations selected from the conceptual domain can be reduced.

The limits associated with the functional form among a set of concepts may be assessed by comparing the original formulation with alternative relations and determining which more accurately describes the set of finding. For instance, some researchers studying fear-arousing communications have postulated a linear relation between level of fear in the message and the amount of attitude change. An alternative formulation is an inverted U function relating the level of fear to the amount of attitude change. If the inverted U function more adequately describes the relations among the concepts, the limits associated with the linear function have been identified, i.e., the relationship between level of fear and attitude change is linear only under certain conditions. This finding reduces uncertainty because the researcher is better able to specify the conditions under which the particular functional forms will hold.

The limits associated with a conceptual model may be assessed by competitively testing this model against alternative formulations. For instance, Brinberg (1979) compared the Fishbein and Triandis model of decision making in their ability to predict a person's intention. The components of Fishbein's model were not sufficient to predict the particular intention (i.e., church attendance) suggesting some limits to the Fishbein theory. For instance, a researcher might infer from this finding that the

Triandis model is more effective in predicting decisions that have a moral component whereas the Fishbein model is more useful for "reasoned" actions. The limits identified by this model comparison do reduce the uncertainty associated with both Fishbein's and Triandis's models because the researcher can specify conditions under which each model will and will not predict better than some alternative.

In sum, the explanatory validity of a research finding can be assessed by conducting robustness analyses and boundary search; that is, by trying to establish the scope and limits of the findings. This assessment can be done with many aspects of the conceptual domain, including the functional relations among the concepts and the sufficiency of the conceptual model.

Summary and Some Implications

The goal of research in any scientific discipline is to acquire knowledge by the reduction of uncertainty in relation to the discipline's "findings." A Validity Network Schema, introduced by Brinberg & McGrath (1982) and elaborated and extended here, describes research as a three stage process, whose outcome is some structured combinations of concepts, methods, and substantive events. In Stage 1, researchers develop, clarify, and select elements and relations from the three basic domains -- conceptual, methodological, and substantive. For this stage, the concept of validity takes on the meaning of value; only those elements and relations are used that are perceived by the researcher to be consistent with the prevailing norms of that scientific community and, therefore, as being worthwhile/useful/appropriate for use.

Stage 2, the research study proper, involves two steps. In step 1, the researcher combines elements and relations from two of the three domains to form an intermediate structure; in step 2, the researcher integrates that structure with the elements and relations from the third domain. With three domains, there are three paths by which this two step process can be carried out. For this stage, research takes on the meaning of correspondence or fit, i.e., the extent to which elements and relations from each of the domains correspond when combined.

Stage 3 of the research process involves the reduction of uncertainty associated with the findings of Stage 2 by exploring the range of respondents, events, contexts, methods, and concepts over which the Stage 2 results do and do not hold. It involves replication to determine whether the finding can be repeated when all conditions are intended to be the same. It also involves the exploration of the robustness and boundaries of the findings -- that is, an attempt to establish the scope and limits of those findings -- with respect to various facets of each of the three domains. In Stage 3, validity takes on the meaning of robustness and generalizability.

Our hope is that this paper will have contributed to the growing literature on validity issues and the research process in several specific ways. First, we hope this presentation of the Validity Network Schema, somewhat more elaborated with respect to Stage 3, will provide a useful framework for considering the very many terms and concepts within the general topic of validity. Second, we hope we have made the point, clearly and convincingly, that efforts to establish the external validity of a set of research findings always and necessarily involves both a search for the range/scope over which the findings do hold and a search for the boundary/

limits of conditions beyond which the findings do not hold. Finally, we hope we have intrigued some readers with some of our suggestions -- so that they will spend some time and thought on them and improve and extend what we have presented here.

References

Anderson, N.H. "Cognitive Algebra." In L. Berkowitz (Ed.) Advances in Experimental Social Psychology, (Vol. 7). New York: Academic Press, 1974.

Ayer, A. Logical Positivism. New York: Free Press, 1959.

Birnbaum, M.H. & B.A. Mellers. "Stimulus Recognition May Mediate Exposure Effects." Journal of Personality and Social Psychology, 1979, 37, 391-394.

Blalock, H.M. Conceptualization and Measurement in the Social Sciences. Beverly Hills, Cal.: Sage Publications, 1982.

Boring, E.G. "A history of introspection." Psychological Review, 1953, 50, 169-189.

Brinberg, D. "An Examination of the Determinants of Intention and Behavior: a Comparison of Two Models." Journal of Applied Social Psychology, 1979, 9, 560-575.

Brinberg, D. & J. E. McGrath. "A Network of Validity Concepts Within the Research Process." In D. Brinberg & L. Kidder (Eds.) New Directions for Methodology of Social and Behavioral Science - Forms of Validity in Research, No. 12. San Francisco: Jossey-Bass, Inc., 1982.

Calder, B.J., L.W. Phillips & A.M. Tybout. "Designing Research for Application." Journal of Consumer Research, 1981, 8, 197-207.

Campbell, D.T. "Social Attitudes and Other Acquired Behavioral Disposition." In S. Koch (Ed.) Psychology: A Study of Science (Vol. 6) New York: McGraw-Hill, 1963.

Campbell, D.T. Comment: Another perspective on a scholarly career: In M.B. Brewer & B.E. Collins, (Eds.) Scientific Inquiry and the Social Sciences. San Francisco: Jossey-Bass Publishers, 1981.

Campbell, D.T. & D.W. Fiske. "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix." Psychological Bulletin, 1959, 56, 81-105.

Campbell, D.T. & E.J. O'Connell. "Methods as Diluting Trait Relationships Rather Than Adding Irrelevant Systematic Variance." In D. Brinberg & L. Kidder (Eds.) New Directions for Methodology of Social & Behavioral Sciences - Forms of Validity in Research, No. 12. San Francisco: Jossey-Bass, Inc., 1982.

Campbell, D.T. & J.C. Stanley. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand-McNally, 1966.

Cook, T.S. & D.T. Campbell Design and Analysis of Quasi-Experiments for Field Settings. Chicago: Rand-McNally, 1979.

Cronbach, L.J. "Beyond the Two Disciplines of Scientific Psychology." American Psychologist, 1975, 30, 116-127.

Cronbach, L.J. Designing Evaluation of Educational Programs. San Francisco: Jossey-Bass, Inc., 1982.

Crowne, D.P. & D. Marlowe. The Approval Motive. New York: Wiley, 1964.

Davis, J.H. Group Performance. Reading, Mass.: Addison-Wesley, 1969

Endler, N.S. & D. Magnusson, (Eds.) Interactional Psychology and Personality. Washington, D.C.: Hemisphere Publishing, 1976.

Engel, J.F. & R.D. Blackwell. Consumer Behavior. New York: The Dryden Press, 1982.

Feyerabend, P.K. "Against Method: Outline of an Anarchistic Theory of Knowledge." In M. Radner & S. Winoker (Eds.) Minnesota Studies in the Philosophy of Science, Vol. 4. Minneapolis: University of Minnesota Press, 1970.

Fishbein, M. "A Theory of Reasoned Action: Some Applications and Implications." In H. Howe & M. Page (Eds.) Nebraska Symposium on Motivation. Lincoln, Nebraska: University of Nebraska Press, 1979.

Fishbein, M. & I. Ajzen. Belief Attitude, Intention and Behavior: An Introduction to Theory and Research. Reading, Mass.: Addison-Wesley Publishing Co., 1975.

Fiske, D.W. "Convergent-Discriminant Validity in Measurements and Research Strategies." In D. Brinberg & L. Kidder (Eds.) New Directions for Methodology of Social and Behavioral Science - Forms of Validity in Research, No. 12. San Francisco: Jossey-Bass, Inc., 1982.

Garner, W.R., H.W. Hake & C.W. Eriksen. "Operationism and the Concept of Perception." Psychological Review, 1956, 63, 149-159.

Gergen, K.J. "Social Psychology as History." Journal of Personality and Social Psychology, 1973, 26, 309-320.

Greenwald, A.G. "Within-Subject Designs: To Use or Not To Use?" Psychological Bulletin, 1976, 83, 314-320.

Grice, R.G. & J.J. Hunter. "Stimulus Intensity Effects Depend Upon the Type of Experimental Design." Psychological Review, 1964, 71, 247-256.

Hovland, C.I., A.A. Lumsdale & F.D. Sheffield. Experiments on Mass Communication. New York: John Wiley & Sons, Inc., 1949.

Jaccard, J., R. Knox & D. Brinberg. "Prediction of Behavior from Beliefs: An Extension and Test of a Subjective Probability Model." Journal of Personality and Social Psychology, 1979, 37, 1239-1249.

Kahle, L.R. & J.J. Berman. "Attitudes Cause Behaviors: A Cross-Lagged Panel Analysis." Journal of Personality and Social Psychology, 1979, 37, p. 315-321.

Kaplan, A. The Conduct of Inquiry: Methodology for Behavioral Science. New York: Chandler Publishing Co., 1964.

Kuhn, T.J. The Structure of Scientific Revolution. Chicago: University of Chicago Press, 1970.

Lerner, M.J. & D. Miller. "Just World Research and the Attribution Process: Looking Back and Ahead." Psychological Bulletin, 1978, 85, 1030-1051.

Lynch, J.G. "On the External Validity of Experiments in Consumer Behavior." The Journal of Consumer Research, 1982, 9, 225-239.

McDougall, W. "An Introduction to Social Psychology." London: Morrison & Gibb, Ltd., 1908.

McGrath, J.E., J. Martin & R.A. Kulka. Judgement Calls in Research. Beverly Hills, Cal.: Sage Publications, Inc., 1982.

Nisbett, R.E. & T.D. Wilson. "Telling More Than We Can Know: Verbal Reports on Mental Processes." Psychological Review, 1977, 84, 231-259.

Popper, K. The Logic of Scientific Discovery. New York: Basic Books, 1959.

Rosenberg, M.J. "The Condition and Consequences of Evaluation Apprehension." In R. Rosenthal & R. Rosnow (Eds.) Artifact in Behavioral Research. New York: Academic Press, 1969.

Runkel, P.J. & J.E. McGrath. Research on Human Behavior: A Systematic Guide to Method. New York: Holt, Rinehart and Winston, Inc., 1972.

Sjoberg, G. & R. Nett. A Methodology for Social Research. New York: Harper and Row, 1968.

Steiner, I. "Whatever Happened to the Group in Social Psychology." Journal of Experimental Social Psychology, 1974, 10, 93-108.

Wimsatt, W.C. "Robustness, Reliability and Overdetermination." In M. Brewer and B. Collins (Eds.) Scientific Inquiry and the Social Sciences. San Francisco: Jossey-Bass, Inc., 1981.

Footnotes

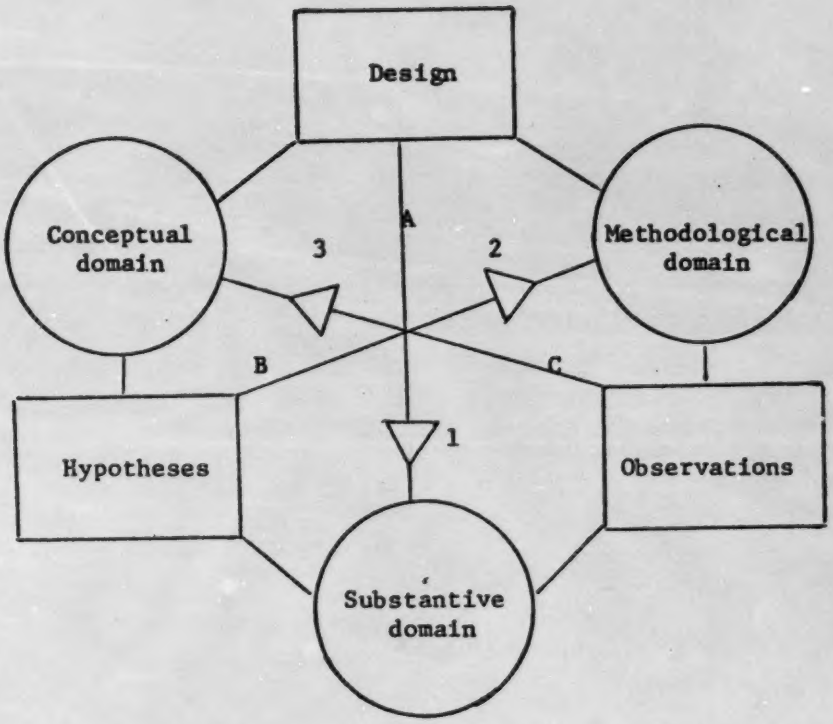
1. The term researcher is used to refer to a particular scientific community. We do not mean to imply that a single researcher need conduct all types of research. For a discipline to reduce the uncertainty associated with a finding, however, the issues raised in this paper need to be considered.

Table 1

Summary of the Facets of Potential Sources of Ambiguity for a Research Finding

Substantive	Methodological	Conceptual
Respondents	Measures	Functional form among concepts
Events - Behaviors	Designs	Sufficiency of model
Contexts - Spatial - Temporal	Strategies	

Figure 1



A. Implementing a design by using it on a set of substantive events
B. Testing a set of hypotheses by evaluating it with an appropriate set of methods
C. Explaining a set of observations by construing it in terms of a set of concepts

1. Ecological Validity
2. Methodological Validity
3. Explanatory Validity